

Solution Techniques for the Stokes System: A priori and a posteriori modifications, resilient algorithms

Markus Huber, Lorenz John, Petra Pustejovska, Ulrich Rüde,
Christian Waluga, Barbara Wohlmuth*

Abstract. This article proposes modifications to standard low order finite element approximations of the Stokes system with the goal of improving both the approximation quality and the parallel algebraic solution process. Different from standard finite element techniques, we do not modify or enrich the approximation spaces but modify the operator itself to ensure fundamental physical properties such as mass and energy conservation. Special local a priori correction techniques at re-entrant corners lead to an improved representation of the energy in the discrete system and can suppress the global pollution effect. Local mass conservation can be achieved by an a posteriori correction to the finite element flux. This avoids artifacts in coupled multi-physics transport problems. Finally, hardware failures in large supercomputers may lead to a loss of data in solution subdomains. Within parallel multigrid, this can be compensated by the accelerated solution of local subproblems. These resilient algorithms will gain importance on future extreme scale computing systems.

Mathematics Subject Classification (2000). Primary 00A05; Secondary 00B10.

Keywords. Stokes system, stabilization, energy-corrected finite element methods, optimal order a priori estimates, fault tolerance, resilience, pollution effect, re-entrant corner, mass conservation

1. Introduction

Advances in numerical methods together with progress in computer technology enables the modeling and simulation of ever more complex physical phenomena with increasingly higher accuracy. Cost aware numerical simulation, embedded in the wider field of computational science and engineering has thus grown to be a fundamental pillar for all quantitative sciences.

The past decades have seen the development of an extensive theoretical foundation of finite element (FE) approximation and for optimal solution algorithms. In the context of simulation technology, the success of this research must be measured by which accuracy can be achieved at which cost. This naturally leads to the question how both accuracy and cost can be quantified. However, an answer to

*This work was supported by the German Research Foundation (DFG) through the Priority Programme 1648 Software for Exascale Computing (SPPEXA).

these seemingly trivial and obvious questions is more difficult than it may appear at first sight. In particular, the cost metrics have undergone a dramatic change due to the advent of ever more complex, heterogeneous and hierarchical parallel computer systems. For the development of numerical methods that will be used in large scale simulations, it has become important to acknowledge that non-parallel computers have ceased to exist. Equally important, we will illustrate that even low cost computers can solve huge problems, provided the models, algorithms, and software have been developed to exploit their full power.

In this article, we will present new research results that explore the duality of accuracy versus cost. We will use the Stokes system for viscous incompressible flow as a guiding example. The Stokes problem in the polyhedral domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, with homogeneous Dirichlet boundary conditions is given by

$$\begin{aligned} -\Delta \mathbf{u} + \nabla p &= \mathbf{f} && \text{in } \Omega, \\ \operatorname{div} \mathbf{u} &= 0 && \text{in } \Omega, \\ \mathbf{u} &= \mathbf{0} && \text{on } \partial\Omega, \end{aligned} \tag{1}$$

where \mathbf{u} denotes the velocity field, p the pressure, and \mathbf{f} represents a given force term.

The methods proposed below are corrections that compensate known numerical deficiencies with minimal additional cost. In the case of re-entrant corners, the local singularity leads to a global deterioration of accuracy. Rather than enriching the approximation space to better capture the singular behavior, we propose a strictly local a priori modification of the energy inner product, showing that once the energy is correctly represented, the global error pollution will be eliminated. This avoids an enrichment of the FE approximation space by, e.g., local mesh refinement that would lead to more complex data structures, a more involved load balancing strategy and an increased communication in solution algorithms. The energy correction, in contrast, requires only the change of a few coefficients of the stiffness matrix, and thus does not lead to a significant increase in cost.

Analogously, we will present an a posteriori correction technique that equips simple stabilized equal order elements with the much-desired local mass conservation. This again avoids the use of more complex discretizations that would in turn lead to an increased cost in the solver and the parallel communication.

The solvers presented here rely on the multigrid principle combined suitably with Krylov space acceleration. We will demonstrate that these methods do not only have optimal complexity, but that they can be implemented with excellent parallel performance. Here we go beyond the classical notion of scalability which is only a necessary but no sufficient prerequisite for a parallel solver being fast. Having a locally conservative scheme and a fast solver allow us to do large-scale long-term simulation runs that arise from simplified Earth mantle convection problems.

On future supercomputers, a fail-safe operation of the complete system may not be guaranteed. In this case, some components may fail and the corresponding part of the running computation may be lost. Then, it is attractive to use algorithms that can compensate for a partial loss of data. This algorithmically based fault

tolerance is an intrinsic correction technique that compensates for large local errors that are naturally introduced when a processor fails.

The rest of this article is structured as follows: In Sect. 2, we present the standard notation and recall the stabilized equal-order finite element discretization. Then, in Sect. 3, we introduce the method of energy-corrections and demonstrate its effectiveness for simple examples in two dimensions. Parallel multigrid solvers for large scale computations and the new fault tolerant algorithms are presented in Sect. 4. Finally, we discuss local mass-conservative correction schemes in Sect. 5 including application to problems arising from geophysics.

2. Preliminaries and finite element discretization

Through this work, we assume Ω to be an open, bounded, simply connected polyhedral domain. We employ standard notation, i.e., $H^k(\Omega)$, $\|\cdot\|_k$, $k \in \mathbb{N}$, denote the Sobolev spaces of functions having square-integrable generalized derivatives up to order k , and the standard Sobolev norm, respectively. The case of $k = 0$ coincides with the Lebesgue space $L^2(\Omega)$. Additionally, we denote the $H^1(\Omega)$ -space with vanishing trace on $\partial\Omega$ by $H_0^1(\Omega)$, and the space of square integrable functions with vanishing mean by $L_0^2(\Omega)$. We also write X^d for a space of vector valued functions which components belong to the function space X .

The main focus of this paper are continuous low equal-order velocity and pressure discretizations. For conforming and shape-regular partitions \mathcal{T}_h of Ω into simplicial elements, we assume finite element spaces $\mathbf{V}_h = V_h^d \cap H_0^1(\Omega)^d$ and $Q_h = V_h \cap L_0^2(\Omega)$, where $V_h := \{v_h \in H^1(\Omega) : v_h|_T \in P_1(T) \ \forall T \in \mathcal{T}_h\}$. It is well-known, that such a pair violates the discrete inf-sup condition. Thus, we work in the following, unless not further specified, with a stabilized finite element scheme. Namely, let us define $a(\mathbf{u}, \mathbf{v}) := \int_{\Omega} \nabla \mathbf{u} : \nabla \mathbf{v} \, dx$ and $b(\mathbf{v}, q) := - \int_{\Omega} q \operatorname{div} \mathbf{v} \, dx$, $\mathbf{v}, \mathbf{u} \in H^1(\Omega)$, $q \in L^2(\Omega)$, for which the discrete variational formulation of (1) reads: Find $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times Q_h$ such that

$$\begin{aligned} a(\mathbf{u}_h, \mathbf{v}_h) + b(\mathbf{u}_h, p_h) &= \langle \mathbf{f}, \mathbf{v}_h \rangle \quad \forall \mathbf{v}_h \in \mathbf{V}_h, \\ b(\mathbf{u}_h, q_h) - c(q_h, p_h) &= \langle f_{\text{st}}, q_h \rangle \quad \forall q_h \in Q_h, \end{aligned} \tag{2}$$

where the additional stabilization terms $c(\cdot, \cdot)$ and $\langle f_{\text{st}}, \cdot \rangle$ are specified by the particular stabilization method. In many cases, we shall assume the standard pressure stabilization for which $\langle f_{\text{st}}, \cdot \rangle = 0$ and

$$c(q_h, p_h) := \sum_{T \in \mathcal{T}_h} c_T(p_h, q_h), \quad c_T(p_h, q_h) := \delta_T h_T^2 \int_T \nabla p_h \cdot \nabla q_h \, dx,$$

where $\delta_T > 0$ is a certain stabilization parameter specified later, or the PSPG stabilization where additionally $\langle f_{\text{st}}, q_h \rangle := - \sum_{T \in \mathcal{T}_h} \delta_T h_T^2 \int_T \mathbf{f} \cdot \nabla q_h \, dx$. For details, see, e.g., [8, 10].

Throughout the paper, for simplicity of notation, we use the symbols \lesssim and \gtrsim for $\leq \text{const.}$ and $\geq \text{const.}$, respectively, where const. represents a generic positive constant independent of the mesh size h .

3. A priori energy correction

In the presence of re-entrant corners, standard numerical schemes on uniformly refined meshes suffer from a pronounced deterioration of the global convergence rates. Here, we focus on the recovery of the optimal convergence rates by a local a priori modification of a standard finite element scheme in two dimensions. For simplicity, we restrict ourselves to the case of a single re-entrant corner. Before we present our method, let us state here the necessary regularity results for the solution of (1), together with its main properties. For that, some weighted Lebesgue spaces with weight as a distance function r to the non-convex domain-vertex¹ x_0 , see, e.g., [1, 26, 28], play an important role. Namely, we define for $\beta \in \mathbb{R}$ and $r(x) = \|x - x_0\|_{l^2}$:

$$L^2_\beta(\Omega) := \left\{ f \in L^1_{\text{loc}}(\Omega) : \int_\Omega |r^\beta f|^2 dx < \infty \right\}, \quad \|f\|_{0,\beta}^2 := \left(\int_\Omega |r^\beta f|^2 dx \right).$$

The weighted Sobolev space of Kondratiev-type of order k , are then given by $H^k_\beta(\Omega) := \left\{ f \in L^2_\beta(\Omega) : r^{\beta-k+|\alpha|} D^\alpha f \in L^2(\Omega), |\alpha| \leq k \right\}$, where $\alpha = (\alpha_1, \alpha_2)$ is a multi-index with $|\alpha| = \alpha_1 + \alpha_2$, and D^α represents the α -th generalized derivative, see, e.g., [26]. The space $H^k_\beta(\Omega)$ is equipped with the norm $\|f\|_{k,\beta}^2 := \sum_{|\alpha| \leq k} \|r^{\beta-k+|\alpha|} D^\alpha f\|_0^2$. The spaces $L^2_\beta(\Omega)$ and $H^k_\beta(\Omega)$ are Hilbert spaces.

3.1. Singularities for Stokes flow at re-entrant corners. A standard spectral analysis applied on the Dirichlet system (1) leads to a specification of the eigenvalues and (generalized) eigenvectors of the Stokes operator that can be used for a local additive composition of the solution, see, e.g., [11, 27, 29] and the original work [26]. In particular for the Dirichlet problem, the eigenvalues are specified by the roots of

$$\lambda_i^2 \sin^2(\omega) - \sin^2(\lambda_i \omega) = 0, \quad \lambda_i \neq 0, \quad \lambda_i \in \mathbb{C}. \quad (3)$$

Note, that the solutions of (3) may be complex and possibly have higher multiplicity, at most two, see Fig. 1.

This is significant, since these roots constrain the regularity of the particular singular functions. The eigenvectors are then, in the polar coordinate-system (r, θ) , log-polynomial functions in r multiplied with smooth functions in θ . Explicitly, for each λ_i with geometric and algebraic multiplicity I_i and $\kappa_{i,j}$, respectively, the singular functions for the velocity and pressure part of the solution are of the form

$$\mathbf{s}_{i,j,k}^u = r^{\lambda_i} \sum_{l=0}^k (\ln r)^l \boldsymbol{\varphi}_{j,k-l}^{(i)}(\theta), \quad s_{i,j,k}^p = r^{\lambda_i-1} \sum_{l=0}^k (\ln r)^l \xi_{j,k-l}^{(i)}(\theta), \quad (4)$$

$k \in \{0, \kappa_{i,j} - 1\}$. Under the assumption that the right-hand side has a higher regularity, i.e., $\mathbf{f} \in L^2_{-\alpha}(\Omega)^2$ for some $\alpha > 1 - \lambda$, together with the assumption

¹In what follows, such a point we simply call singular point.

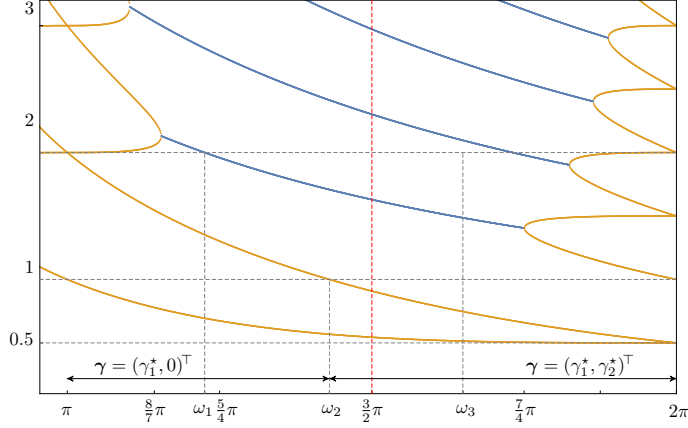


Figure 1. Distribution of the real part of the eigenvalues λ with respect to the angle ω of the re-entrant corner. The blue lines represent complex eigenvalues, the orange ones represent the real eigenvalues. The points of bifurcations, also the point where $\lambda_2 = 1$, have increased multiplicity. In the horizontal axis, ω_i , $i = 1, 2, 3$, represent the following cases: ω_1 is defined such that $\lambda_1 + \lambda_2 = 2$, i.e., $\omega_1 = 1.22552\dots\pi$, ω_2 is the unique angle such that $\lambda_2 = 1$, i.e., $\omega_2 = 1.430296\dots\pi$, and ω_3 represents the angle for which $\lambda_1 + \text{Re}(\lambda_3) = 2$, i.e., $\omega_3 = 1.64941\dots\pi$. In the plot, there are also included two ω -intervals, (π, ω_2) and $(\omega_2, 2\pi)$, see Theorem 3.1 and the discussion after.

that λ_i are simple roots for all $i \in [1, N]$, there holds the following expansion of the solution near the singular point (for simplicity we skip the j, k indices):

$$\mathbf{u} = \sum_{i=1}^N c_i \mathbf{s}_i^u + \mathbf{U}, \quad p = \sum_{i=1}^N c_i s_i^p + P, \quad (5)$$

where $(\mathbf{U}, P) \in H_{-\alpha}^2(\Omega)^2 \times H_{-\alpha}^1(\Omega)$ is called a smooth remainder. For our needs, we will require a decomposition with two singular functions, i.e., $N = 2$. The assumption on simple multiplicity of the eigenvalues is in that case violated only for one angle ω_2 in the whole $(\pi, 2\pi)$ range.

It is clear from the structure of the singular functions (4) that the integrability and boundedness of the gradients is limited which results in a reduced regularity of the solution in domains with re-entrant corners. On the other hand, the special polynomial structure in r of the singularities can be well-handled in the framework of the Kondratiev-type Sobolev spaces. Hence, one can obtain a generalized regularity result

$$\|\mathbf{u}\|_{2,\beta} + \|p\|_{1,\beta} \lesssim \|\mathbf{f}\|_{0,\beta}, \quad (6)$$

see, e.g., [21]. The explicit form of the singular parts of the solution (\mathbf{u}, p) , decomposition (5), a priori bounds following from the structure of the singularities and the smooth remainder, together with regularity (6) are then used in the formula-

tion and the proof of the convergence behavior of the energy-corrected methods, as they are stated below.

3.2. Pollution effect. It is well-known that the influence of the non-convex corners significantly reduces the global accuracy of standard numerical methods. More precisely, standard Galerkin approximations on a sequence of uniformly refined meshes of the Stokes problem are, in general, of the order:

$$\|\nabla(\mathbf{u} - \mathbf{u}_h)\|_0 + h^{-\lambda_1} \|\mathbf{u} - \mathbf{u}_h\|_0 + \|p - p_h\|_0 = \mathcal{O}(h^{\lambda_1}), \quad (7)$$

cf. [2, 6, 17]. This global effect, which is commonly referred to as *pollution* cannot be cured by considering different error norms. Also, it is exhibited by all standard (piecewise polynomial) approximations, independently of the approximation order. For the illustration of the pollution effect, we present the errors and convergence rates for the stable Taylor–Hood element in Tab. 1, where we compare a typical global pollution in the standard approximation with the non-polluted interpolation; cf. also Fig. 2 for some illustration.

$\ \mathbf{u} - \mathbf{u}_h\ _0$	eoc	$\ p - p_h\ _0$	eoc	$\ \mathbf{u} - \mathbf{u}_h\ _{0,\alpha}$	eoc	$\ p - p_h\ _{0,\alpha}$	eoc
6.24297e-02	–	3.76799e-01	–	2.94081e-02	–	4.73148e-01	–
3.03536e-02	1.04	2.00709e-01	0.91	9.38624e-03	1.65	1.72502e-01	1.46
1.20753e-02	1.33	1.27266e-01	0.66	3.21263e-03	1.55	7.58580e-02	1.19
4.90503e-03	1.30	8.06119e-02	0.66	1.36660e-03	1.23	3.50865e-02	1.11
2.06841e-03	1.25	5.23079e-02	0.62	6.20441e-04	1.14	1.65028e-02	1.09
9.04139e-04	1.19	3.47209e-02	0.59	2.88024e-04	1.11	7.78711e-03	1.08
4.06348e-04	1.15	2.33942e-02	0.57	1.34837e-04	1.09	3.67251e-03	1.08

Table 1. Errors and the convergence rates of the standard Galerkin approximation in standard and weighted norms ($\alpha = 1 - \lambda_1$) in the case of the L-shape domain (see Fig. 3 third from left) using the (higher-order) Taylor–Hood element, computed for an exact solution $(\mathbf{s}_1^u + \mathbf{s}_2^u, s_1^p + s_2^p - |\Omega|^{-1} \langle s_1^p + s_2^p, 1 \rangle_\Omega)$. Estimated convergence rates of the errors in standard norms are for the velocity: $2\lambda_1 = 1.09$ and the pressure: $\lambda_1 = 0.54$.

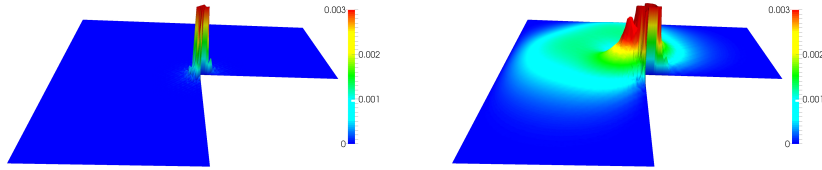


Figure 2. Illustration of the global propagation of the pollution in a L-shape domain with exact solution: $(\mathbf{s}_1^u, s_1^p - |\Omega|^{-1} \langle s_1^p, 1 \rangle_\Omega)$ for the Taylor–Hood element; left: non-polluted $|\mathbf{s}_1^u - I_h^2 \mathbf{s}_1^u|$, right: $|\mathbf{s}_1^u - \mathbf{s}_{1,h}^u|$ on level $\ell = 4$.

3.3. The energy-corrected scheme. Let us now focus on the a priori energy-modification for the stabilized scheme. We closely follow the theory given

in [24], thus it serves as a reference for additional information or rigorous proofs for the presented statements. In view of this approach, we define for a given correction vector $\gamma = (\gamma_1, \gamma_2)^\top$, at this point not closely specified, a mesh-dependent bilinear form $a_h(\cdot, \cdot)$ for $\mathbf{v}_h, \mathbf{w}_h \in \mathbf{V}_h$ by:

$$\begin{aligned} a_h(\mathbf{w}_h, \mathbf{v}_h) &:= a(\mathbf{w}_h, \mathbf{v}_h) - d_h(\mathbf{w}_h, \mathbf{v}_h) \\ &= \sum_{T \in \mathcal{T}_h \setminus (\mathcal{L}_h^1 \cup \mathcal{L}_h^2)} a_T(\mathbf{w}_h, \mathbf{v}_h) + \sum_{i=1,2} (1 - \gamma_i) \sum_{T \in \mathcal{L}_h^i} a_T(\mathbf{w}_h, \mathbf{v}_h), \end{aligned} \quad (8)$$

where $a_T(\mathbf{w}_h, \mathbf{v}_h) := \int_T \nabla \mathbf{w}_h : \nabla \mathbf{v}_h dx$, \mathcal{L}_h^i are defined as sets of elements which are in the i -th element layer around the singular point, and a stabilization term

$$c_h(q_h, z_h) := \sum_{T \in \mathcal{T}_h \setminus (\mathcal{L}_h^1 \cup \mathcal{L}_h^2)} c_T(q_h, z_h) + \sum_{i=1,2} \frac{1}{1 - \gamma_i} \sum_{T \in \mathcal{L}_h^i} c_T(q_h, z_h). \quad (9)$$

The stabilized energy-corrected finite element approximation to (2) is then formulated for the bilinear forms $a_h(\cdot, \cdot)$ and $c_h(\cdot, \cdot)$. Namely, we find $(\mathbf{u}_h^m, p_h^m) \in \mathbf{V}_h \times Q_h$ such that

$$\begin{aligned} a_h(\mathbf{u}_h^m, \mathbf{v}_h) + b(\mathbf{v}_h, p_h^m) &= \langle \mathbf{f}, \mathbf{v}_h \rangle_\Omega \quad \forall \mathbf{v}_h \in \mathbf{V}_h, \\ b(\mathbf{u}_h^m, q_h) - c_h(q_h, p_h^m) &= \langle f_{\text{st}}, q_h \rangle_\Omega \quad \forall q_h \in Q_h. \end{aligned} \quad (10)$$

By definition (8), and (9), the bilinear forms $a_h(\cdot, \cdot)$ and $c_h(\cdot, \cdot)$ differ from the ones in the standard stabilized Galerkin approximation only on an $\mathcal{O}(h)$ -neighborhood of the singular point.

To measure the impact of the pollution in the L^2 -approximation error, we define the energy defect function (also called pollution function) of the discrete scheme as:

$$g_h(\mathbf{u}, p) := a(\mathbf{u} - \mathbf{u}_h^m, \mathbf{u} - \mathbf{u}_h^m) + 2b(\mathbf{u} - \mathbf{u}_h^m, p - p_h^m) - d_h(\mathbf{u}_h^m, \mathbf{u}_h^m) - c_h(p_h^m, p_h^m). \quad (11)$$

Note, that the pollution function g_h , as defined in (11), can be expressed in the case of a stable finite element formulation of the incompressible homogeneous Dirichlet problem by:

$$g_h(\mathbf{u}, p) = a(\mathbf{u}, \mathbf{u}) - a_h(\mathbf{u}_h^m, \mathbf{u}_h^m),$$

hence the name energy defect function. Also, since the correction will be constructed in such a way that the pollution $g_h(\mathbf{u}, p)$ is controlled, we speak about energy-corrected finite element schemes.

3.4. Theoretical estimates. Let us present here the main result from [24], which was proven for stable linear approximations (more precisely for the P_1 -iso P_2/P_1 element). If explicitly mentioned, we additionally require a local symmetry in angular direction of the mesh around the re-entrant corner; this condition is denoted by (Sh). Note also, that although the main result is formulated for maximal interior angles in the range of $\omega \in (\pi, \omega_3)$, the performance of the modification was shown numerically for all $\omega \in (\pi, 2\pi)$. One can verify that

$$\|\mathbf{u} - \mathbf{u}_h^m\|_{0,\alpha} = \mathcal{O}(h^2) \iff g_h(\mathbf{u}, p) = \mathcal{O}(h^2),$$

and thus, the following theorem represents the sufficient condition for the optimal convergence. The necessary condition, i.e., the implication from left- to right-hand side, is rather trivial to prove and free of additional assumptions.

Theorem 3.1 (Optimal order convergence). *Let $\omega \in (\pi, \omega_3)$, $\omega \neq \omega_2$, $\mathbf{f} \in L^2_{-\alpha}(\Omega)^2$ for some α such that $\alpha > 1 - \lambda_1$, let (\mathbf{V}_h, Q_h) be a stable pair with $c_h(\cdot, \cdot) = 0$, and, additionally, let (Sh) hold if $\omega \in [\omega_1, \omega_3)$. Moreover, let $\gamma \in \mathbb{R}^2$ be such that*

$$g_h(\mathbf{s}_i^u, s_i^p) = \mathcal{O}(h^2), \quad \text{for } \begin{cases} i = 1 & \text{if } \omega \in (\pi, \omega_2), \\ i = 1, 2 & \text{if } \omega \in (\omega_2, \omega_3), \end{cases} \quad (12)$$

where ω_1 , ω_2 and ω_3 are defined as in Fig 1. Then, the modified Galerkin approximation $(\mathbf{u}_h^m, p_h^m) \in \mathbf{V}_h \times Q_h$ converges optimally, i.e.,

$$\|\mathbf{u} - \mathbf{u}_h^m\|_{0,\alpha} \lesssim h^2 \|\mathbf{f}\|_{0,-\alpha}, \quad \text{and} \quad \|\mathbf{u} - \mathbf{u}_h^m\|_{1,\alpha} + \|p - p_h^m\|_{0,\alpha} \lesssim h \|\mathbf{f}\|_{0,-\alpha} \quad (13)$$

First of all, let us describe the assumptions in Theorem 3.1 and the proof-technique, after we shall discuss its generalization to the stabilized case.

At the beginning, we would like to point out here, that $\gamma = (\gamma_1, \gamma_2)^\top \in \mathbb{R}^2$ is constrained such that the bilinear forms $a_h(\cdot, \cdot)$ and $c_h(\cdot, \cdot)$ remain uniformly elliptic and continuous, respectively. This means, since the element layers are disjoint, we actually require $\gamma \in \mathcal{B}_R^\infty(0)$ ($\mathcal{B}_R^\infty(0)$ being a ball with center in the origin of radius $R < 1$ in the l^∞ norm). Also, if the first condition of (12) has to be satisfied, i.e., $\omega \in (\pi, \omega_2)$, we represent the vector γ for simplicity of notation by $\gamma = (\gamma_1, 0)^\top$. In that case, we speak about a single- (or one-) parameter correction modification. Such a case largely inherits the properties of the modification as derived in [12] for the Laplace equation.

The proof of (13) is based on duality arguments and strongly relies on the linearity of the system. Under the assumption, that the right-hand side is of increased regularity, the decomposition (5) can be used on the exact and the approximative solutions (note that the eigenvectors (\mathbf{s}_i^u, s_i^p) are solutions of the Stokes equations themselves), and thus, the pollution in the system can be split into several parts, namely pollution caused by the singular functions, terms corresponding to the smooth remainder and the so-called cross-terms. By standard duality arguments, a uniform inf-sup condition with respect to weighted norms, a priori bounds of the smooth remainder in negatively weighted norms, as introduced by [6], one can show that the terms with smooth remainder are of second order, and thus they do not pollute. The same holds for the cross-terms. However, the proof is technically more involved. More precisely, since the orthogonality of the eigenvectors is, in general, not preserved by the finite element approximations, we have to apply generalized Wahlbin arguments and use the local symmetry condition (Sh) on the mesh. Note here, that the drop of (Sh) for small angles comes from the sufficient regularity of the second singular function, and thus the symmetry arguments do not have to be used. The contribution to the pollution g_h by the singular functions cannot be neglected, therefore it is enforced by the modification itself. This exactly reflects the condition (12). As one can see, for angles smaller as ω_2 we have for the second singular function: $(\mathbf{s}_2^u, s_2^p) \in H^2_{-\varepsilon}(\Omega)^2 \times H^1_{-\varepsilon}(\Omega)$, and thus its pollution is

again a priori of second order. This means, it does not have to be corrected. The correction vector γ is then reduced to a single scalar value, as noted above.

In Theorem 3.1, the weight α is bounded from below which originates in the assumption on the decomposition (5). Note, that also for the results discussed in [24], some additional assumptions on the upper bound are specified. These are mainly for technical reasons, namely restricting the number of singular functions in the decomposition and to ensure the validity of some auxiliary lemmas.

Let us now turn our attention to the extension of Theorem 3.1 to stabilized finite elements. Our approach is motivated by the equivalence of the MINI and a stabilized P_1 - P_1 element. For that, we define a space of element-wise bubble functions:

$$\mathbf{B}_h = \{v_h \in \mathcal{C}(\overline{\Omega}) : v_h|_T \in \text{span}\{\lambda_1^T, \lambda_2^T, \lambda_3^T\} \ \forall T \in \mathcal{T}_h\}^2,$$

where λ_i^T , $i = 1, \dots, 3$, denote the barycentric coordinates on the element T , see also [2]. The MINI element is then defined via the product $(\mathbf{V}_h \oplus \mathbf{B}_h) \times Q_h$. The equivalence of the MINI element and the stabilized one, as defined above, (see [9] for the standard case) holds also for the case of the energy corrected scheme if $c_h(\cdot, \cdot)$ is selected appropriately. Namely, we have the following.

Theorem 3.2. *Let \mathbf{f} be piecewise constant on each element $T \in \mathcal{T}_h$. Then the energy corrected MINI element discretization is equivalent to the energy corrected P_1 - P_1 discretization with PSPG stabilization (10) where the stabilization parameter $\delta_T = \frac{c_{1,T} c_2}{\alpha_T}$. Further, the element-wise constants are determined by*

$$\langle \phi_T, 1 \rangle_T = c_{1,T} h_T^2, \quad \langle \phi_T, 1 \rangle_T = c_2 |T|, \quad \langle \nabla \phi_T, \nabla \phi_T \rangle_T = \alpha_T,$$

where $\phi_T \in B_h|_T$ denotes the bubble function, scaled to a maximal value of one, on each element $T \in \mathcal{T}_h$, and $c_2 = \frac{9}{20}$.

Proof. Let us consider the energy corrected MINI finite element formulation, and denote the solution by $\mathbf{u}_h = \mathbf{u}_h^l + \mathbf{u}_h^b$, where $\mathbf{u}_h^l \in \mathbf{V}_h$ and $\mathbf{u}_h^b = \sum_{T \in \mathcal{T}_h} \mathbf{u}_{b,T} \phi_T \in \mathbf{B}_h$, with $\mathbf{u}_{b,T} \in \mathbb{R}^2$. Note that the linear and bubble functions are A -orthogonal, i.e., $a(\mathbf{v}_h^l, \phi_T \mathbf{e}_i) = 0$, $i = 1, 2$, for all $\mathbf{v}_h^l \in \mathbf{V}_h^1$ and $\phi_T \in B_h$. Let us consider the test function $\mathbf{v}_h = \phi_T \mathbf{e}_i$, $i = 1, 2$, on T , then for the first equation of the variational formulation (10), we obtain that

$$(1 - \gamma_T) a(\phi_T \mathbf{u}_{b,T}, \phi_T \mathbf{e}_i) + b(\phi_T \mathbf{e}_i, p_h) = \langle \mathbf{f}, \phi_T \mathbf{e}_i \rangle_T,$$

where γ_T represents the correction parameter on the individual element. Note, $\gamma_T = 0$ for $T \in \mathcal{T}_h \setminus (\mathcal{L}_h^1 \cup \mathcal{L}_h^2)$. Using integration by parts and the fact that $\phi_T = 0$ on ∂T we have:

$$(\mathbf{u}_{b,T} \cdot \mathbf{e}_i)(1 - \gamma_T) \langle \nabla \phi_T, \nabla \phi_T \rangle_T = \langle \mathbf{f} - \nabla p_h, \phi_T \mathbf{e}_i \rangle_T.$$

Then, applying the assumption of a piecewise constant \mathbf{f} , we get that

$$\mathbf{u}_{b,T} = \frac{1}{\alpha_T(1 - \gamma_T)} (\mathbf{f} - \nabla p_h)|_T \langle \phi_T, 1 \rangle_T.$$

Using this, the bilinear form $b(\cdot, \cdot)$ can be rewritten as:

$$\begin{aligned} b(\phi_T \mathbf{u}_{b,T}, q_h) &= (\mathbf{u}_{b,T} \cdot \nabla q_h)|_T \langle \phi_T, 1 \rangle_T = \frac{1}{\alpha_T(1 - \gamma_T)} (\mathbf{f} - \nabla p_h) \cdot \nabla q_h|_T \langle \phi_T, 1 \rangle_T^2 \\ &= \frac{c_{1,T} c_2}{\alpha_T(1 - \gamma_T)} h_T^2 \langle \mathbf{f} - \nabla p_h, \nabla q_h \rangle_T, \end{aligned}$$

and thus

$$b(\mathbf{u}_h^l, q_h) - \sum_{T \in \mathcal{T}_h} \frac{1}{1 - \gamma_T} \delta_T h_T^2 \langle \nabla p_h, \nabla q_h \rangle_T = - \sum_{T \in \mathcal{T}_h} \frac{1}{1 - \gamma_T} \delta_T h_T^2 \langle \mathbf{f}, \nabla q_h \rangle_T,$$

which corresponds to the PSPG stabilization. \square

In the case that the computational domain is the L-shape domain with a mesh as depicted by the third one from the left in Fig. 3, the stabilization parameter δ_T is independent of the element and equal to $1/160$. The equivalence from Lemma 3.2 also motivates the extension to other stabilizations, like the Dohrman–Bochev approach [7] or the local projection stabilization [15]. In that case we proceed as before, as we scale the bilinear form $c_h(\cdot, \cdot)$ on each element $T \in \mathcal{L}_h^i$ by the term $(1 - \gamma_i)^{-1}$, for $i = 1, 2$. We note that this is the inverse scaling of the bilinear form $a_h(\cdot, \cdot)$.

3.5. Identification of the correction vector γ . Up to this point, we assumed that the vector γ , chosen such that $g_h(\mathbf{u}, p) = \mathcal{O}(h^2)$, is known and exists. If this is the case, then the practical realization into an existing code is straightforward, since it only requires a simple re-scaling of a few local stiffness matrices.

The numerical determination of the unknown γ was first proposed for the Laplace equation in [12] and theoretically studied in [30]. It is constructed as a limit-vector in a component-wise meaning (or a limit-value), of a sequence $\{\gamma_h\}_{h>0}$ solving

$$g_h(\mathbf{s}_i^u, s_i^p) = \mathbf{0},$$

for $i = 1$ if $\omega \in (\pi, \omega_2)$, and $i = 1, 2$ if $\omega \in (\omega_2, 2\pi)$. We understand the notation above in a following functional-setting sense: For each $\gamma \in \mathcal{B}_R^\infty(0)$, the approximation operators $R^u : H_0^1(\Omega)^2 \rightarrow \mathbf{V}_h$ and $R^p : L_0^2(\Omega) \rightarrow Q_h$ are considered as functions of γ . Thus, on each computational level, corresponding to a certain $h > 0$, we seek a root γ_h of

$$a(\mathbf{s}_i^u, \mathbf{s}_i^u) - [a_h(R^u(\gamma)\mathbf{s}_i^u, R^u(\gamma)\mathbf{s}_i^u) + c_h(R^p(\gamma)s_i^p, R^p(\gamma)s_i^p)],$$

again, i from the same set as above. The existence and uniqueness of the sequence $\{\gamma_h\}_{h>0}$ can be proved with the help of a fixed point theorem. Scaling arguments and the properties of the energy correction function yield the convergence of γ_h , and we denote the asymptotic value by star, i.e., $\gamma^* = (\gamma_1^*, \gamma_2^*)^\top =$

$(\lim_{h \rightarrow 0^+} \gamma_{1,h}, \lim_{h \rightarrow 0^+} \gamma_{2,h})^\top$. Then one can show that for $\gamma = \gamma^*$ the assumption (12) is satisfied. To do so one has to construct a suitable subsequence and exploit the properties of the Jacobi matrix of g_h . The computation of γ^* can be accelerated by a combination of a nested hierarchical Newton scheme with some extrapolation technique on a coarse element patch.

We remark that γ^* depends only on the interior angle ω , on the local element patch and the finite element type, but not on the mesh-size nor on the rest of the mesh/domain. Moreover, many re-entrant corners can be handled by local modifications each of which carrying its own correction parameter.

3.6. Numerical examples. In this section, we consider several numerical examples for the stabilized schemes (2) and (10). The correction parameters are estimated as described above, see Tab. 2, for the initial grids as depicted in Fig. 3, i.e., for the scenarios where $\omega \in \{8/7\pi, 5/4\pi, 3/2\pi, 7/4\pi\}$. We always consider a computational setting for which $(\mathbf{s}_1^u + \mathbf{s}_2^u, s_1^p + s_2^p - |\Omega|^{-1} \langle s_1^p + s_2^p, 1 \rangle_\Omega)$ is the exact solution. Note, that for $\omega = 8/7\pi$ we do not require the local symmetry of the mesh at the singular point, for both $\omega = 8/7\pi, 5/4\pi$ a single-parameter correction can be used, while $\omega = 3/2\pi, 7/4\pi$ cases require two correction parameters. The last case $\omega = 7/4\pi$ represents the extension of the derived theory for angles bigger than ω_3 . For simplicity, we represent only the results for the velocity, see Tab.3, more precisely, we show the weighted $L_\alpha^2(\Omega)$ error convergences and their rates.

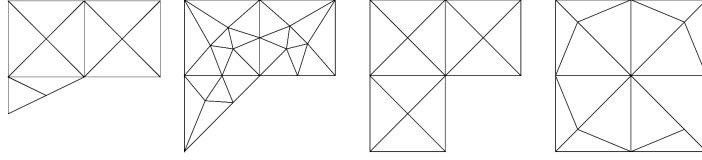


Figure 3. Initial grids for several types of the re-entrant corner, namely $\omega \in \{8/7\pi, 5/4\pi, 3/2\pi, 7/4\pi\}$. All the geometries are bounded by a unit square, particularly $\Omega_{3/2\pi} = (-1, 1)^2 \setminus ([0, 1] \times [-1, 0])$.

ω	γ^* (single correction)	ω	γ^* (two corrections)
$8/7\pi$	$(0.011364, 0.0)^\top$	$3/2\pi$	$(0.161708, -0.183984)^\top$
$5/4\pi$	$(0.017676, 0.0)^\top$	$7/4\pi$	$(-0.483852, 0.273889)^\top$

Table 2. Estimated asymptotic γ^* for the geometries and local element patches of Fig. 3.

To demonstrate the convergence of the pressure part of the solution, we additionally include Tab.4 for $\omega = 3/2\pi$, presenting also the rates and errors in the case of the standard $L^2(\Omega)$ -norm. For the velocity, we obtain in that case the best approximation order, namely $\lambda_1 + 1 = 1.54$, instead of the polluted rate $2\lambda_1 = 1.09$, see (7). In the case of the pressure-error, we have increased the weight to $\alpha_p = \alpha + 0.5$, to restore its superconvergence behavior.

Finally, we return to the motivational problem of Fig. 2, i.e., a higher order

$\omega = 8/7\pi$		$\omega = 5/4\pi$		$\omega = 3/2\pi$		$\omega = 7/4\pi$	
$\ \mathbf{u} - \mathbf{u}_h^m\ _{0,\alpha}$	eoc	$\ \mathbf{u} - \mathbf{u}_h^m\ _{0,\alpha}$	eoc	$\ \mathbf{u} - \mathbf{u}_h^m\ _{0,\alpha}$	eoc	$\ \mathbf{u} - \mathbf{u}_h^m\ _{0,\alpha}$	eoc
6.80763e-02	–	6.43853e-02	–	9.87423e-02	–	6.31650e-02	–
2.21098e-02	1.62	2.36951e-02	1.44	3.76712e-02	1.39	1.79778e-02	1.81
6.06479e-03	1.87	6.78762e-03	1.80	1.28515e-02	1.55	4.55198e-03	1.98
1.59057e-03	1.93	1.81517e-03	1.90	3.48892e-03	1.88	1.06693e-03	2.09
4.07657e-04	1.96	4.68415e-04	1.95	8.87484e-04	1.97	2.47722e-04	2.11
1.03206e-04	1.98	1.18907e-04	1.98	2.19857e-04	2.01	5.82728e-05	2.09
2.59642e-05	1.99	2.99543e-05	1.99	5.37256e-05	2.03	1.41046e-05	2.05

Table 3. Errors and the convergence rates of the velocity approximation measured in the weighted norm with $\alpha = 1 - \lambda_1$.

standard norm				weighted norm			
$\ \mathbf{u} - \mathbf{u}_h^m\ _0$	eoc	$\ p - p_h^m\ _0$	eoc	$\ \mathbf{u} - \mathbf{u}_h^m\ _{0,\alpha}$	eoc	$\ p - p_h^m\ _{0,\alpha_p}$	eoc
4.70326e-02	–	7.92268e-01	–	9.87423e-02	–	4.03685e-01	–
3.87065e-02	2.81	4.61252e-01	0.78	3.76712e-02	1.39	2.04358e-01	0.98
1.88110e-02	1.04	2.30692e-01	0.99	1.28515e-02	1.55	5.03504e-02	2.02
7.07091e-03	1.41	1.25586e-01	0.87	3.48892e-03	1.88	1.83549e-02	1.46
2.51514e-03	1.49	6.98617e-02	0.84	8.87484e-04	1.97	6.74790e-03	1.44
8.78492e-04	1.51	3.97820e-02	0.81	2.19857e-04	2.01	2.37574e-03	1.50
3.04097e-04	1.53	2.33040e-02	0.77	5.37256e-05	2.03	8.20768e-04	1.53

Table 4. Convergence results of the velocity and pressure modified approximation measured in the standard and weighted norms, with minimal admissible value $\alpha = 1 - \lambda_1$ and increased weight for pressure $\alpha_p = \alpha + 0.5$ to restore the superconvergence, in the case of L-shape domain.

approximation, and illustrate on the left of Fig. 4 the reduction of the pollution by the energy-corrected scheme, as derived above, for the same computational setting. On the right, we show the solution of an optimal control problem on a domain with multiple (50) re-entrant corners. Here at each corner, the modified scheme can be applied locally with the same correction parameters as for the L-shape domain. Many quantities of interest such as stress intensity factors, eigenvalues or the flow rate can then accurately be determined on uniformly refined meshes, and no mesh grading techniques are required.

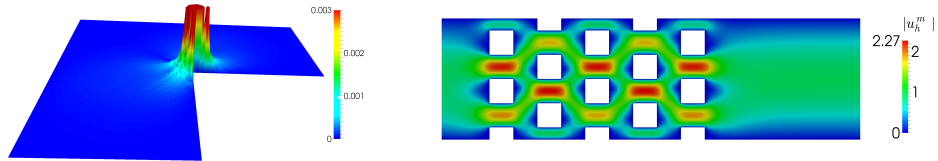


Figure 4. Left: Reduced pollution effect for the modified Taylor-Hood scheme (compare with Fig. 2); Namely we plot $|\mathbf{u} - \mathbf{u}_h^m|$. Right: Multiple re-entrant corners.

4. Multigrid and fault tolerant algorithms

In this section, we consider the realization of efficient multigrid solvers and fault tolerant algorithms for the Stokes system. Our implementation employs the hierarchical hybrid grids (HHG) finite element framework, see, e.g., [5, 4]. HHG combines the flexibility of unstructured finite element meshes with the performance advantage of structured grids in a block-structured approach. It shows excellent scalability up to the largest supercomputers available [18] and reaches parallel textbook multigrid efficiency [19].

In the following, we denote by $n_u = \dim \mathbf{V}_h$ and $n_p = \dim Q_h$ the number of degrees of freedom for velocity and pressure, respectively. Then, the isomorphisms $\mathbf{u}_h \leftrightarrow \mathbf{u} \in \mathbb{R}^{n_u}$ and $p_h \leftrightarrow \mathbf{p} \in \mathbb{R}^{n_p}$ are obvious. The discrete variational formulation (2) is then equivalent to the linear system

$$\mathcal{K} \begin{pmatrix} \mathbf{u} \\ \mathbf{p} \end{pmatrix} := \begin{pmatrix} A & B^\top \\ B & -C \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{f}_{\text{st}} \end{pmatrix}, \quad (14)$$

where $\mathcal{K} \in \mathbb{R}^{(n_u+n_p) \times (n_u+n_p)}$. The matrix A has a 3×3 diagonal block structure consisting of three discrete representations of the Laplace operator.

In the following, we describe a multigrid solver for this structure and a multigrid based fault tolerant solution strategy.

4.1. Multigrid solver. Iterative methods for solving the saddle point system can e.g. be constructed based on the pressure Schur complement (SCG) or on a preconditioned MINRES method, see, e.g., [13, 32].

Here, we consider an *all-at-once* multigrid method for the linear system (14). The key point in a multigrid method for the entire saddle point problem is the construction of a suitable smoother. We shall consider Uzawa type smoothers, see, e.g., [3, 31, 34]. This idea is based on a preconditioned Richardson method, where in iteration $k+1$ the system for $(\mathbf{u}_{k+1}, \mathbf{p}_{k+1})$ must be solved

$$\hat{\mathcal{K}} \begin{pmatrix} \mathbf{u}_{k+1} - \mathbf{u}_k \\ \mathbf{p}_{k+1} - \mathbf{p}_k \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{f}_{\text{st}} \end{pmatrix} - \mathcal{K} \begin{pmatrix} \mathbf{u}_k \\ \mathbf{p}_k \end{pmatrix}. \quad (15)$$

Here \mathcal{K} denotes the system matrix (14) of the Stokes equations and $\hat{\mathcal{K}}$ an appropriate preconditioner. The corresponding iteration matrix is then given by $\mathcal{M} := I - \hat{\mathcal{K}}^{-1}\mathcal{K}$. We note that for the convergence of (15), the analysis of \mathcal{M}^k is needed, while for showing a qualitative smoothing property, the matrix polynomial $(I - \mathcal{M})\mathcal{M}^k$ must be analyzed [25, 31]. Approaches towards a full quantitative multigrid analysis based on Fourier techniques are presented in [16]. In general several choices for $\hat{\mathcal{K}}$ are available, we restrict ourselves to

$$\hat{\mathcal{K}} := \begin{pmatrix} \hat{A} & 0 \\ B & -\hat{S} \end{pmatrix},$$

where \hat{A} and \hat{S} are suitable smoothers for the upper left block A and the pressure Schur complement $S := BA^{-1}B^\top + C$, respectively. The application of $\hat{\mathcal{K}}$ results

in the algorithm of the inexact Uzawa method, where we smooth the velocity part in a first step and in a second step the pressure, i.e.

1. $\mathbf{u}_{k+1} = \mathbf{u}_k + \hat{A}^{-1}(\mathbf{f} - A\mathbf{u}_k - B^\top \mathbf{p}_k),$
2. $\mathbf{p}_{k+1} = \mathbf{p}_k + \hat{S}^{-1}(B\mathbf{u}_{k+1} - C\mathbf{p}_k - \mathbf{f}_{\text{st}}).$

Other variants of Uzawa-type multigrid methods, such as the symmetric version, are for instance considered in [31]. In each smoothing step, we consider for the velocity smoother \hat{A} , a combination of a forward and backward communication reducing variant of a Gau -Seidel scheme with an additional row-wise red-black coloring in each of the parallel distributed subdomains. Within the parallel data structures this results in a combined block Jacobi- and symmetric Gau -Seidel updating scheme. For the pressure, we use a SOR applied on the matrix C with the under-relaxation parameter $\omega = 0.3$. These smoothers are then applied within a multigrid V-cycle, where we vary the number of smoothing steps from 3 pre- and 3 post smoothing steps on the finest level to 5 pre- and 5 post-smoothing steps on coarser levels. This leads to a good efficiency.

As a solver on the coarsest level, several choices are available. This includes parallel sparse direct solvers, such as PARDISO [22], or algebraic multigrid methods, such as hypre [14]. For this performance study, we avoid the dependency on external libraries and employ a preconditioned MINRES as coarse-grid solver for the saddle point system (14). The preconditioner consists of CG and a lumped mass matrix M . The relative accuracy is set to $\varepsilon = 5 \cdot 10^{-3}$.

Numerical examples. In the following, we present several numerical examples, illustrating the robustness and scalability of our solver. As a computational domain, we consider the unit cube $\Omega = (0, 1)^3$ with an initial mesh \mathcal{T}_{-2} consisting of 6 tetrahedra. The coarse grid \mathcal{T}_0 is a twice refined initial grid, where each tetrahedron is decomposed into 8 new ones. The right-hand side $\mathbf{f} = \mathbf{0}$ and for the initial guess $\mathbf{x}_0 = (\mathbf{u}_0, \mathbf{p}_0)^\top$, we distinguish between velocity and pressure part. Here \mathbf{u}_0 is a random vector with values in $[0, 1]$, while the initial random pressure \mathbf{p}_0 is scaled by h^{-1} , i.e., with values in $[0, h^{-1}]$. This is important in order to represent a less regular pressure, since the velocity is generally considered as a $H^1(\Omega)$ and the pressure as a $L^2(\Omega)$ function. Unless further specified, we use the relative reduction of the residual by $\varepsilon = 10^{-8}$ as a stopping criteria. Numerical examples are performed on a low cost desktop machine, an Intel Xeon CPU E2-1226 v3, 3.30GHz and 32 GB shared memory for serial computations and on JUQUEEN (J lich Supercomputing Center, Germany) for massively parallel computations.

In Tab. 5, we present results for several refinement levels L and a fixed coarse grid \mathcal{T}_0 . We observe that the multigrid convergence rates are independent of the mesh size, as expected. We also include results for the Schur complement CG method [18], where we observe that the Uzawa multigrid solver is typically a factor of two faster and is more memory efficient. The number of operator evaluations for the individual blocks, which are computed by $n_{\text{op}} = \sum_{\ell=0}^L 8^{\ell-L} n_{\text{op},\ell}$, where $n_{\text{op},\ell} \in \mathbb{N}$ denotes the number of operator evaluations on level ℓ , are presented too. These numbers are also reflected by the time to solution of the individual solvers.

L	DoF	SCG		Uzawa	
		iter	TtS [s]	iter	TtS [s]
2	$1.4 \cdot 10^4$	26	0.11	9	0.08
3	$1.2 \cdot 10^5$	28	0.56	8	0.29
4	$1.0 \cdot 10^6$	28	3.33	8	1.79
5	$8.2 \cdot 10^6$	28	24.28	8	12.70
6	$6.6 \cdot 10^7$	31	205.84	8	95.85
7	$5.3 \cdot 10^8$	out of mem.		8	730.77

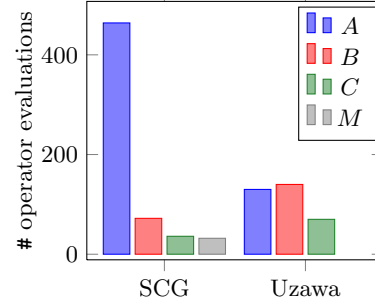


Table 5. Iteration numbers and time-to-solution (TtS) (left) and number of operator evaluations for $L = 6$ (right) for the serial case.

Nodes	Threads	DoF	iter	TtS [s]	T_w [s]	T_c [%]
1	24	$6.6 \cdot 10^7$	7 (1)	65.71	58.80	10.5
6	192	$5.3 \cdot 10^8$	7 (3)	80.42	64.40	19.9
48	1 536	$4.3 \cdot 10^9$	7 (5)	85.28	65.03	23.7
384	12 288	$3.4 \cdot 10^{10}$	7 (10)	93.19	64.96	30.3
3 072	98 304	$2.7 \cdot 10^{11}$	7 (20)	114.36	66.08	42.2
24 576	786 432	$2.2 \cdot 10^{12}$	8 (40)	177.69	78.24	56.0

Table 6. Weak scaling result times for the Uzawa multigrid method; number of iterations, including number of iterations of the preconditioner on the coarse grid, time-to-solution (TtS), TtS without the time spent by the coarse grid iterations (T_w), and the time spent by the coarse grid in percents (T_c).

Scaling results are present for up to 786 432 threads and more than 10^{12} degrees of freedom (DoF). We use 32 threads per node and 6 tetrahedra in the initial mesh per thread. The finest computational level is obtained by a 7 times refined initial mesh. Excellent scalability for the multilevel part of the algorithm is achieved, see Tab. 6. We point out that the coarse grid solver in these examples considered as a stand-alone solver lacks optimal complexity. However, even for the largest example with more than 10^{12} DoF it still requires less than 60% of the overall time.

4.2. Fault tolerant algorithms. In the future era of exa-scale computing systems, highly scalable implementations will execute up to billions of parallel threads on millions of compute nodes. In this scenario, fault tolerance will become a necessary property of hardware, software and algorithms. Nevertheless, nowadays commonly used redundancy approaches, e.g., check-pointing, will be too costly, due to the memory and energy consumption. Our approach, on the other hand, incorporates the resilience strategy directly into the multigrid solver, and thus does not require additional memory.

In [23], we introduce a methodology and data-structure to efficiently recover

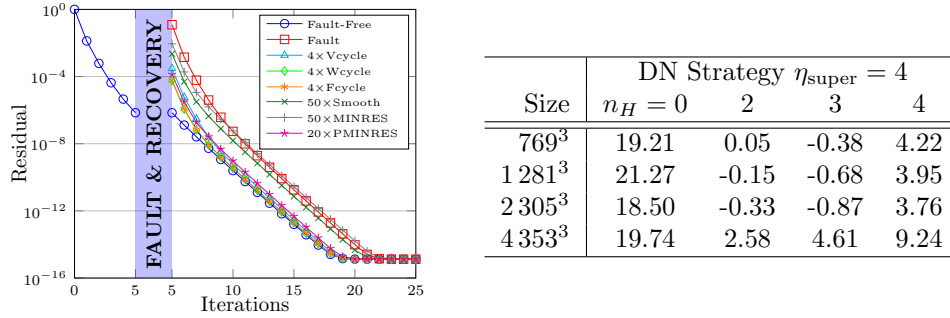


Figure 5. Left: Convergence of the relative residual for different local recovery strategies, Right: Global recovery in terms of a domain partitioning concept with over-balancing by a factor of $\eta_{\text{super}} = 4$ in the case of the Laplace operator.

lost data due to a processor crash (hard fault) when solving elliptic PDEs with multigrid algorithms. We consider a fault model where a processor stores the mesh data of a subdomain including all its refined levels in the multigrid hierarchy. Therefore, when a processor crashes all data is lost in the faulty domain $\Omega_F \subset \Omega$. The healthy domain $\Omega_H \subset \Omega$ is unaffected by the fault, and data in this domain remains available. Further, we denote by $\Gamma := \partial\Omega_F \cap \partial\Omega_H$ the interface. The nodes associated with Γ are used to communicate between neighboring processors by introducing *ghost copies* that redundantly exist on different processors. Therefore, a complete recovery of these nodes is possible without additional redundant storage.

In order to recover the lost nodal values $(\mathbf{u}_F, \mathbf{p}_F)$ in Ω_F , we propose to solve a local subproblem in Ω_F of (1) with Dirichlet boundary conditions on Γ for velocity and pressure, respectively. To guarantee that the local system is uniformly well-posed, we formally include a compatibility condition obtained from the normal components of the velocity. If the local solution in Ω_F is computed while the global process is halted, then this is a local recovery strategy. For a first analysis, we assume that the local recovery is free of cost, i.e., that it can be executed in zero time.

In [23] this local strategy is extended to become a global recovery strategy. To this end, the solution algorithm proceeds asynchronously in the faulty and the healthy domain such that no process remains idle. Temporarily, the two subdomains are decoupled at the interface Γ , and the recovery process is accelerated by delegating more compute resources to it. This acceleration is termed the *superman strategy*. Once the recovery has proceeded far enough and has caught up with the regular solution process, both subdomains are re-coupled and the regular global iteration is resumed. These approaches result in a time- and energy-efficient recovery of the lost data.

Numerical examples. In Fig. 5 (left), we consider a test scenario in $\Omega = (0, 1)^3$ in which we continuously apply multigrid V(3,3)-cycles of the Uzawa multigrid

method introduced in Sect. 4.1. In total 23 iterations are needed to reach the round-off limit of 10^{-15} . During the iteration, a fault is provoked after 5 iterations. This affects 2.1% of the pressure and velocity unknowns. We continue with global solves after the recovery by solving the local auxiliary problem iteratively in Ω_F . As approximate subdomain solvers we compare the fine grid Uzawa-smoother, the minimal residual method, the block-diagonal preconditioned MINRES (PMINRES) method and V, F, W-cycles in the variable smoothing variant from Sect. 4.1. For the block-preconditioner in case of PMINRES a standard V-cycle and a lumped mass-matrix M are used. Fig. 5 (left) displays also the cases in which no fault appears (fault-free) and when no recovery is performed after the fault.

After the fault, we observe that the residual jumps up and when no recovery is performed, the iteration must start almost from the beginning. A higher pre-asymptotic convergence rate after the fault helps to catch up, so that only four additional iterations are required. This delay can be further reduced by a local recovery computation, but only local multigrid cycles are found to be efficient recovery methods. Four local recovery multigrid cycles are sufficient such that the round-off limit is reached with the same number of iterations as in the fault-free case. The other recovery methods either yield no significant improvement or too many recovery iterations would be required.

The table on the right of Fig. 5 summarizes the performance of the global recovery in terms of the time delay (in seconds compute time) as compared to an iteration without faults. The tests are performed for a large Laplace problem discretized with up to almost 10^{11} unknowns. The undisturbed solution takes 50.49 seconds with 14 743 cores on JUQUEEN. Two faults are provoked, one after 5 V-cycles, one after 9. Both faults are treated with both the global recovery strategy and a local *superman* process that is $\eta_{\text{super}} = 4$ times as fast as a regular processor. Faulty and healthy domain remain decoupled for n_H cycles with the Dirichlet-Neumann strategy, i.e., solving a Dirichlet problem on Ω_F and a Neumann problem on Ω_H , then the regular iteration is resumed. The case $n_H = 0$ corresponds to performing no recovery at all and leads to a delay of about 20 seconds. By the superman recovery and global re-coupling after $n_H = 2$ cycles, the delay can be reduced to just a few seconds. In some cases the fault-affected computation is even faster than the regular one, as indicated by negative time delays in the table.

5. A posteriori mass correction

A common problem of continuous pressure finite element methods, such as the stabilized scheme considered in this work, is that they do not preserve the physical concept of mass-conservation in a local sense. Although the discrete solution $\mathbf{u}_h \in \mathbf{V}_h$ satisfies a weak form of the mass-balance equation, it can, in general, not be considered as element-wise mass conservative. Due to the lack of element-piecewise constants in the pressure space, we cannot expect a property of the form

$$\int_{\partial T} \mathbf{u}_h \cdot \mathbf{n} \, ds = 0, \quad T \in \mathcal{T}_h,$$

to hold, which is a crucial prerequisite to enable an element-by-element post-processing of the discrete solution to obtain strongly divergence-free velocities. However, we will demonstrate how a local conservation property can be obtained also for continuous pressure interpolations. For this we need to change the viewpoint which requires some additional notation that is introduced next.

5.1. Dual mesh. Based on the triangulation \mathcal{T}_h , which we refer to \mathcal{T}_h as the *primal* mesh, we construct a *dual* mesh as it is often done for finite-volume methods. To be precise, let $\mathcal{T}_i \subseteq \mathcal{T}_h$ denote the nodal patch associated with the vertex \mathbf{x}_i , $i = 1, \dots, N$, of the primal mesh, and let $\lambda_i^T(\mathbf{x})$ denote the barycentric coordinate of $\mathbf{x} \in \mathbb{R}^d$ with respect to the vertex \mathbf{x}_i belonging to any element $T \in \mathcal{T}_i$. We define

$$B_i^T := \{\mathbf{x} : \lambda_i^T(\mathbf{x}) \geq \lambda_j^T(\mathbf{x}), i \neq j\},$$

and we set $B_i := \bigcup_{T \in \mathcal{T}_i} B_i^T$. This construction results in a partition of Ω into d -polytopes B_i which we shall refer to as the set $\mathcal{B}_h := \{B_i, i = 1, \dots, N\}$ of *control volumes*. For an illustration of typical control volumes in $d = 2, 3$ dimensions, we refer to Fig. 6. The boundary ∂B_i of every such volume B_i can be decomposed

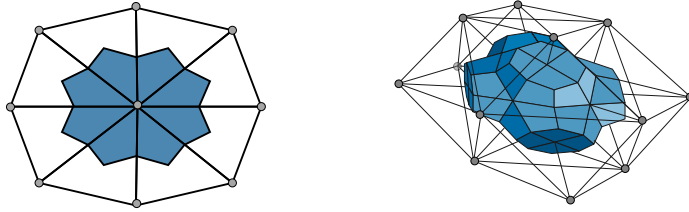


Figure 6. Illustration of a nodal patch \mathcal{T}_i around a vertex \mathbf{x}_i with the associated dual control volume B_i (shaded) for $d = 2$ (left) and $d = 3$ (right).

into the set of internal control facets $\gamma_{ij}^T := \partial B_i \cap \partial B_j \cap T : B_j \neq B_i$ and boundary facets $\gamma_{i\partial}^T := \partial \Omega \cap \partial B_i \cap \partial T$, where B_i and $B_j \in \mathcal{B}_h$ are the control volumes defined with respect to the two vertices \mathbf{x}_i and \mathbf{x}_j , which are both end points of an edge of some $T \in \mathcal{T}_h$. With each γ_{ij}^T we associate a unique normal \mathbf{n}_{ij}^T oriented from B_i^T towards B_j^T . We note that the two faces γ_{ij}^T and γ_{ji}^T are geometrically the same but have a differently oriented normal. We are now prepared to discuss mass conservation properties on the dual mesh.

5.2. Corrected mass flux. In the following, we consider a wider class of continuous-pressure discretization schemes. After we state the main result, we discuss simplifications that are possible if the choice of the finite-element spaces is restricted to stabilized linear equal-order elements as discussed in this work. For the moment let us put the previous definitions aside and assume that $\mathbf{V}_h \times Q_h$ is a continuous mixed finite element space which may be stable or unstable. For unstable choices, we require that appropriate stabilization terms are added to the

variational form such that a unique discrete solution exists.

Given a finite element solution $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times Q_h$, we define a mass flux defect

$$\kappa_{ij}^T(\mathbf{u}_h, p_h) := \frac{1}{(d+1)|\gamma_{ij}^T|} (\mathcal{R}_i^T(\mathbf{u}_h, p_h) - \mathcal{R}_j^T(\mathbf{u}_h, p_h)),$$

where the local residual fluxes \mathcal{R}_i^T are defined as

$$\mathcal{R}_i^T(\mathbf{u}_h, p_h) := \int_{B_i^T} \operatorname{div} \mathbf{u}_h \, dx - \int_T \operatorname{div} \mathbf{u}_h \phi_i \, dx - c_T(p_h, \phi_i) - \langle f_{st}, \phi_i \rangle_T,$$

with $\phi_i \in Q_h$ denoting the linear nodal shape function associated with the local vertex i of element $T \in \mathcal{T}_h$. Finally, we define the corrected mass flux $j(\mathbf{u}_h, p_h)$ which is oriented and defined facet-wise as

$$j(\mathbf{u}_h, p_h)|_{\gamma_{ij}^T} := \mathbf{u}_h|_{\gamma_{ij}^T} \cdot \mathbf{n}_{ij}^T - \kappa_{ij}^T(\mathbf{u}_h, p_h).$$

Given these preparations, we formulate our main observation, which is the conservation of the corrected fluxes over the boundary of a dual control volume.

Theorem 5.1. *Let $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times Q_h$ denote the solution of the discrete Stokes problem and assume that $c_T(\cdot, 1) + \langle f_{st}, 1 \rangle_T = 0$ for every $T \in \mathcal{T}_h$. Then the corrected mass-flux $j(\mathbf{u}_h, p_h)$ is locally conservative, i.e.,*

$$\int_{\partial B} j(\mathbf{u}_h, p_h) \, ds = 0, \quad B \in \mathcal{B}_h.$$

The proof is based on summation arguments and the weak mass-conservation property of the discrete solution, see [20, Theorem 3.2].

Revisiting the construction of the mass-correction, we observe that for a stable pairing without additional stabilization terms (such as the classical Taylor-Hood element), the defect $\kappa_{ij}^T(\mathbf{u}_h, p_h)$ obviously does not depend on p_h , while for stabilized equal-order linear elements the correction does not depend on \mathbf{u}_h since we find by a simple integration that the divergence terms cancel inside each element. The remaining terms can often be considerably simplified, e.g., for the stabilized discretization we consider in this paper, we find that the flux correction assumes the simple form

$$\kappa_{ij}^T(p_h) = \delta_T h_T^2 \nabla p_h|_T \cdot \mathbf{n}_{ij}^T.$$

which can be straightforwardly evaluated.

Given a locally conservative mass-flux on the dual mesh, we could next solve local equilibration problems to obtain local fluxes also with respect to the primal mesh. This allows the lifting of the correction into $H(\operatorname{div})$ -conforming finite element spaces and thereby allows us to define a strongly mass-conservative velocity solution with the same order of convergence as the original finite element solution without post-processing. Details are outlined in [20].

However, when coupling the equal-order linear stabilized scheme to transport problems, we can even circumvent the reconstruction step by directly solving the transport equations on the dual mesh in a finite volume fashion, using the corrected mass fluxes for advection.

5.3. Conservative coupling to transport equations. In many applications non-isothermal models (e.g., mantle convection) or reactive models (e.g., the simulation of thrombus formation in the human blood-stream) are considered. What these models have in common is that we couple fluid flow in the form of (1) to a (set of) transport equation(s) of the form:

$$\partial_t \vartheta + \operatorname{div}(\vartheta \mathbf{u} - \varepsilon \nabla \vartheta) = 0 \quad \text{in } \Omega \times (0, t_{\text{end}}],$$

for some simulation end time $t_{\text{end}} > 0$, where ϑ denotes the transported quantity, $\varepsilon > 0$ plays the role of a conductivity, and \mathbf{u} is an often dominant advective velocity field that typically results from a flow simulation. To make the problem well-posed, we consider an initial temperature field $\vartheta(t = 0) = \vartheta^0$ and homogeneous Neumann boundary conditions $\varepsilon \nabla \vartheta \cdot \mathbf{n} = 0$ on the whole boundary $\partial\Omega$ of the computational domain, where \mathbf{n} denotes an outward-pointing unit-normal vector. This is done only for simplicity of the exposition and as we will see in the numerical examples the extension to other types of boundary conditions is possible.

Of course, in non-isothermal models the flow-properties such as the forcing or the viscosity may again depend non-linearly on ϑ . However, for simplicity we shall consider only a one-way coupling between flow and transport here. For more general models with full non-isothermal coupling we refer to our recent work [33].

Our aim is to complement the equal-order linear stabilized scheme with a finite-volume solver that associates nodal degrees of freedom with unknown coefficients. Therefore, we define the following space on the barycentric dual mesh:

$$R_h := \{r_h \in L^2(\Omega) : r_h|_B \in P_0(B), B \in \mathcal{B}_h\},$$

We note that due to the fact that $\dim R_h = \dim V_h = n_p$, we can define a natural bijection between these two spaces. More precisely, given the nodal basis ϕ_i of V_h and the piecewise constant basis χ_i of R_h , we can define the natural transfer operators $\pi_h : \sum_{i=1}^{n_p} q_i \chi_i \rightarrow \sum_{i=1}^{n_p} q_i \phi_i$, and $\bar{\pi}_h : \sum_{i=1}^{n_p} q_i \phi_i \rightarrow \sum_{i=1}^{n_p} q_i \chi_i$ which by construction satisfy the property $\pi_h \bar{\pi}_h = \bar{\pi}_h \pi_h = \text{id}$.

Based on the discrete velocity solution of (2) or (10), the transport equation for the temperature can be then advanced by an upwind finite-volume scheme which we state in form of a Petrov-Galerkin method: given $(\mathbf{u}_h^n, p_h^n) \in \mathbf{V}_h \times Q_h$ and $\vartheta_h^n \in V_h$, find $\vartheta_h^{n+1} \in V_h$ such that

$$\sum_{B_i \in \mathcal{B}_h} \int_{B_i} \bar{\pi}_h \vartheta_h^{n+1} r_h \, dx = \sum_{B_i \in \mathcal{B}_h} \int_{B_i} \vartheta_h^n r_h \, dx - \Delta t_n \mathcal{A}(\vartheta_h^n, r_h),$$

for all $r_h \in R_h$, where the use of $\bar{\pi}_h$ introduces a mass-lumping, and we set the bilinear form

$$\mathcal{A}(\vartheta_h^n, r_h) := \sum_{B_i \in \mathcal{B}_h} \int_{\partial B_i \setminus \partial\Omega} (\mathbf{j}(\mathbf{u}_h^n, p_h^n) \langle \vartheta_h^n \rangle_{\text{up}} - \varepsilon \nabla \vartheta_h^n \cdot \mathbf{n}_i) r_h \, ds,$$

where \mathbf{n}_i is the unit normal vector pointing outward of B_i . To ensure stability, we

use an up-winding, namely

$$\langle \vartheta_h^n \rangle_{\text{up}}|_{\partial B_i \cap \gamma_{ij}^T} := \begin{cases} \vartheta_h^n|_{B_i} & \text{if } J_{ij}^T \geq 0 \\ \vartheta_h^n|_{B_j} & \text{if } J_{ij}^T \leq 0 \end{cases}, \quad i \neq j,$$

where we define the net flux $J_{ij}^T := \int_{\gamma_{ij}^T} \mathbf{j}(\mathbf{u}_h^n, p_h^n)|_{\partial B_i} \, ds$ based on the locally conservative facet flux function

$$\mathbf{j}(\mathbf{u}_h^n, p_h^n)|_{\partial B_i} := (\mathbf{u}_h^n - \delta_T h_T^2 \nabla p_h^n)|_{\partial B_i} \cdot \mathbf{n}_i.$$

The distinctive feature of this method is that it can be implemented with minimal overhead, using only nodal data structures on semi-structured or even fully unstructured simplicial meshes. This collocated data layout makes the scheme particularly interesting for the implementation into highly performant stencil-based implementations and legacy codes. Of course, the construction above assumes that the discrete Stokes problem is solved exactly, which is often not feasible in practice, e.g., when iterative solvers are employed. However, the mass-defect introduced by solver inaccuracies are only related to the iteration error in contrast to the discretization error as it is common for methods that do not obey local conservation properties.

5.4. Numerical examples. In the following we would like to show two examples in which the effect of the mass-corrections can be clearly observed. Firstly, we show a geometry in which flow is accelerated and decelerated, leading to regions with poor mass-conservation. We transport a concentration through this domain and show solutions obtained by the corrected coupling scheme and a non-corrected approach. Secondly, we discuss a more involved example from [33] in which we consider non-isothermal fluid properties and induce a forcing by buoyancy terms.

5.4.1. Pipe with obstacles. As a simple example, we consider a pipe which is cylindrical with unit radius and aligned with the z -axis. Inside we place three spherical obstacles as shown in Fig. 7. We further neglect forcing terms and diffusion, i.e., we set $\mathbf{f} = \mathbf{0}$ and $\varepsilon = 0$. The inflow at the left boundary ($z = 0$) is prescribed as $\mathbf{g} = (0, 0, 1 - x^2 - y^2)^\top$, at the right boundary we consider free outflow conditions, and for the remaining boundary we impose a no-slip condition. The transported quantity ϑ is set to $\vartheta = 1$ at the left boundary and can exit the pipe domain freely at the right boundary.

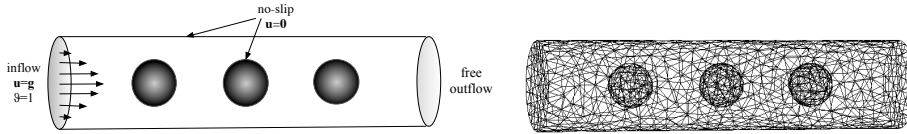


Figure 7. Left: illustration of the geometry and the boundary conditions. Right: coarsest mesh.

We conduct a mesh-refinement study on a series of three meshes with increasing resolution (cf. Fig. 7) in which we choose our time-steps according to the CFL

stability condition. In Fig. 8, we compare our results at $t_{\text{end}} = 10$ for a non-conservative scheme as well as for the same series of experiments conducted with the conservatively coupled scheme. It can be observed that the flux-correction keeps the concentration profiles within physical bounds, and already a visual inspection reveals that it produces physically better solutions with less resolution than the uncorrected method. In the non-conservative case, we observe solution overshoots of around 100% even for the finest mesh. The application of a simple limiter to these overshoots would take mass/energy out of the system, and thus we cannot expect reasonable results when using moderate resolutions with a straightforward coupling. The conservative mass-correction is a simple and elegant way to overcome this deficiency.

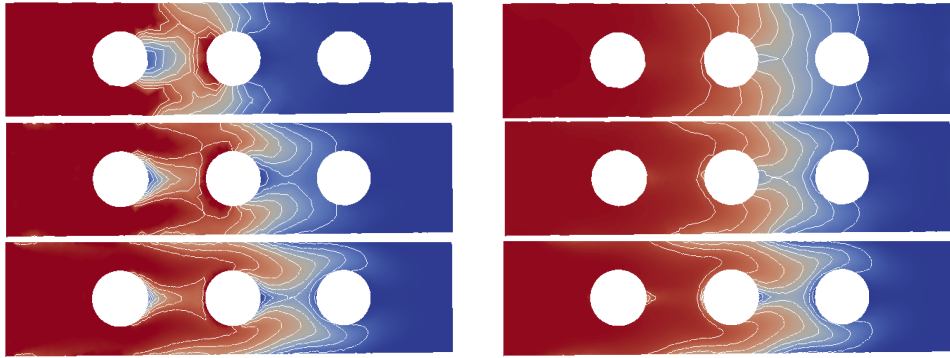


Figure 8. Cross-section in the y/z -plane of the computed solution on mesh-levels 0, 1, and 2 for the uncorrected scheme (left) and the corrected scheme (right); the plotting range is limited to $[0, 1]$ in the left picture.

5.4.2. Non-isoviscous case. In the following, we consider two examples for the non-isoviscous case, both regarding simulations of the Earth mantle. Firstly, the full non-isothermal coupling of buoyancy-driven flow models, where we have a non-linear bidirectional coupling between flow and transport. In particular for this problem, the conservative coupling is crucial in order to avoid a spurious forcing on the velocity field as well as non-physical values of the temperature-dependent viscosity. Secondly, the simulation of the velocity field in the Earth mantle, where the boundary velocities for the Dirichlet condition on the surface and the temperature profile for the right-hand side are given by measured data. On the inner boundary, we use free-slip, i.e., $\mathbf{u} \cdot \mathbf{n} = (\nu D(\mathbf{u}) \mathbf{n} - p \mathbf{n}) \cdot \mathbf{t}_i = 0$ where \mathbf{t}_i , $i = 1, 2$, denote the tangential vectors.

In these cases the momentum equation is given by

$$-\text{div}(\nu D(\mathbf{u})) + \nabla p = \mathbf{f},$$

where $D(\mathbf{u}) = 1/2(\nabla \mathbf{u} + \nabla \mathbf{u}^\top)$, denotes the symmetric part of the velocity gradient and \mathbf{f} a temperature dependent forcing term. We consider in the following the spherical shell $\Omega = \{x \in \mathbb{R}^3 : 0.55 < \|x\|_2 < 1\}$ as a computational domain with

two different viscosities, either temperature depending or piecewise constant

$$\nu_1 = e^{1/2-\vartheta}, \quad \nu_2 = \begin{cases} 1 & \|x\|_2 < 0.936, \\ 0.001 & \text{else.} \end{cases}$$

For both examples the computational mesh is an icosahedral spherical mesh consisting of roughly a hundred million individual tetrahedra. For the first example with ν_1 , we consider thermal convection where the interior temperature at the boundary is set to $\vartheta(r = 0.55) = 1$ and at the surface to $\vartheta(r = 1) = 0$. Initially we prescribe an interpolation between core and surface temperature with a spherical harmonic perturbation. We solve the coupled problem with the previously described SCG multigrid solver for the Stokes part, coupled to an explicit time-integration of the finite-volume transport scheme with flux-correction for 11 500 time-steps on a mid-sized department cluster. A plot of a temperature iso-surface is depicted in Fig. 9 (left). For further details on the setup and the simulation parameters we refer to [33]. For the second example with the discontinuous viscosity ν_2 , where the jump roughly appears in the upper mantle (asthenosphere), we consider as a solver the Uzawa multigrid method, described in Sect. 4.1. A plot of the simulation result is depicted in Fig. 9 (right) which shows the typical behavior of the velocity field and the region of the viscosity jump.

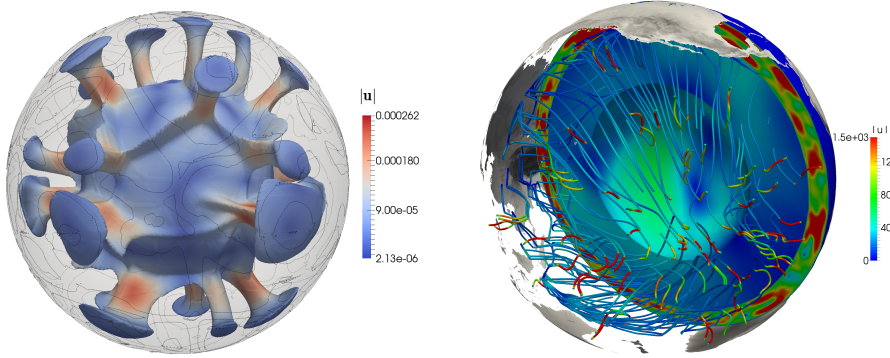


Figure 9. Iso-surfaces for $\vartheta = 0.4$ colored by the velocity magnitude after 11 500 timesteps for ν_1 (left) and magnitude of the velocity \mathbf{u} for the real data simulation, including a viscosity jump ν_2 (right).

Acknowledgements The authors gratefully acknowledge the support by the Gauss Center for Supercomputing (GCS), the Jülich Supercomputing Centre (JSC), and the German Research Foundation (DFG) through the grants WO 671/14-1 and WO 671/11-1.

References

- [1] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces*. Academic Press, New York, London, 2003.

- [2] D. N. Arnold, F. Brezzi, and M. Fortin. A stable finite element for the Stokes equations. *Calcolo*, 21(4):337–344, 1984.
- [3] R. E. Bank, B. D. Welfert, and H. Yserentant. A class of iterative methods for solving saddle point problems. *Numer. Math.*, 56(7):645–666, 1990.
- [4] B. Bergen, T. Gradl, U. Rüde, and F. Hülsemann. A massively parallel multigrid method for finite elements. *Computing in Science and Engineering*, 8(6):56–62, 2006.
- [5] B. Bergen and F. Hülsemann. Hierarchical hybrid grids: data structures and core algorithms for multigrid. *Numer. Linear Algebra Appl.*, 11:279–291, 2004.
- [6] H. Blum. The influence of reentrant corners in the numerical approximation of viscous flow problems. In *Numerical treatment of the Navier-Stokes equations (Kiel, 1989)*, volume 30 of *Notes Numer. Fluid Mech.*, pages 37–46. Vieweg, Braunschweig, 1990.
- [7] P. B. Bochev and C. R. Dohrmann. A stabilized finite element method for the Stokes problem based on polynomial pressure projections. *Internat. J. Numer. Methods Fluids*, 46(2):183–201, 2004.
- [8] D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2013.
- [9] F. Brezzi, M. O. Bristeau, L. P. Franca, M. Mallet, and G. Rogé. A relationship between stabilized finite element methods and the Galerkin method with bubble functions. *Comput. Methods Appl. Mech. Engrg.*, 96(1):117–129, 1992.
- [10] F. Brezzi and J. Douglas, Jr. Stabilized mixed methods for the Stokes problem. *Numer. Math.*, 53(1-2):225–235, 1988.
- [11] M. Dauge. Stationary Stokes and Navier-Stokes systems on two- or three-dimensional domains with corners. I. Linearized equations. *SIAM J. Math. Anal.*, 20(1):74–97, 1989.
- [12] H. Egger, U. Rüde, and B. Wohlmuth. Energy-corrected finite element methods for corner singularities. *SIAM J. Numer. Anal.*, 52(1):171–193, 2014.
- [13] H. C. Elman, D. J. Silvester, and A. J. Wathen. *Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics*. Oxford University Press, New York, 2005.
- [14] R. D. Falgout, J. E. Jones, and U. M. Yang. The design and implementation of hypre, a library of parallel high performance preconditioners. In *Numerical solution of partial differential equations on parallel computers*, volume 51 of *Lect. Notes Comput. Sci. Eng.*, pages 267–294. Springer, Berlin, 2006.
- [15] S. Ganesan, G. Matthies, and L. Tobiska. Local projection stabilization of equal order interpolation applied to the Stokes problem. *Math. Comp.*, 77(264):2039–2060, 2008.
- [16] F. J. Gaspar, Y. Notay, C. W. Oosterlee, and C. Rodrigo. A simple and efficient segregated smoother for the discrete Stokes equations. *SIAM J. Sci. Comput.*, 36(3):A1187–A1206, 2014.
- [17] V. Girault and P. A. Raviart. *Finite Element Methods for Navier-Stokes Equations*. Springer, New York, 1986.
- [18] B. Gmeiner, U. Rüde, H. Stengel, C. Waluga, and B. Wohlmuth. Performance and Scalability of Hierarchical Hybrid Multigrid Solvers for Stokes Systems. *SIAM J. Sci. Comput.*, 37(2):C143–C168, 2015.

- [19] B. Gmeiner, U. Rüde, H. Stengel, C. Waluga, and B. Wohlmuth. Towards textbook efficiency for parallel multigrid. *Numer. Math. Theory Methods Appl.*, 8, 2015.
- [20] B. Gmeiner, C. Waluga, and B. Wohlmuth. Local mass-corrections for continuous pressure approximations of incompressible flow. *SIAM J. Numer. Anal.*, 52(6):2931–2956, 2014.
- [21] B. Guo and C. Schwab. Analytic regularity of Stokes flow on polygonal domains in countably weighted Sobolev spaces. *J. Comput. Appl. Math.*, 190(1-2):487–519, 2006.
- [22] M. Hagemann, O. Schenk, and A. Wächter. Matching-based preprocessing algorithms to the solution of saddle-point problems in large-scale nonconvex interior-point optimization. *Comput. Optim. Appl.*, 36(2-3):321–341, 2007.
- [23] M. Huber, B. Gmeiner, U. Rüde, and B. Wohlmuth. Resilience for mutlgrid software at the extreme scale, 2015. Submitted.
- [24] L. John, P. Pustejovska, U. Rüde, and B. Wohlmuth. Energy-corrected finite element methods for the Stokes system, 2015. In preparation.
- [25] L. John, B. Wohlmuth, and W. Zulehner. On the Smoothing Properties of Block Preconditioners for Saddle Point Problems, 2015. In preperation.
- [26] V. A. Kondratiev. Boundary value problems for elliptic equations in domains with conical or angular points. *Trans. Moscow Math. Soc.*, 16:227–313, 1967.
- [27] V. A. Kozlov, V. G. Mazya, and J. Rossmann. *Spectral Problems Associated with Corner Singularities of Solutions to Elliptic Equations*. Mathematical surveys and monographs. American Mathematical Society, 2001.
- [28] A. Kufner. *Weighted Sobolev spaces*. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1985. Translated from the Czech.
- [29] P. Märkl and A.-M. Sändig. Singularities of the stokes system in polygons. *Preprint-Reihe: Institut für Angewandte Analysis und Numerische Simulation, Universität Stuttgart*, 2008/009, 2008.
- [30] U. Rüde, C. Waluga, and B. Wohlmuth. Nested Newton strategies for energy-corrected finite element methods. *SIAM J. Sci. Comput.*, 36(4):A1359–A1383, 2014.
- [31] J. Schöberl and W. Zulehner. On Schwarz-type smoothers for saddle point problems. *Numer. Math.*, 95(2):377–399, 2003.
- [32] R. Verfürth. A combined conjugate gradient-multigrid algorithm for the numerical solution of the Stokes problem. *IMA J. Numer. Anal.*, 4(4):441–455, 1984.
- [33] C. Waluga, B. Wohlmuth, and U. Rüde. Mass-corrections for the conservative coupling of flow and transport on collocated meshes, 2015. Submitted.
- [34] W. Zulehner. Analysis of iterative methods for saddle point problems: a unified approach. *Math. Comp.*, 71(238):479–505, 2002.

Universität Erlangen–Nürnberg, Cauerstrasse 11, 91058 Erlangen, Germany

E-mail: ulrich.ruede@fau.de

Technische Universität München, Boltzmannstrasse 3, 85748 Garching, Germany

E-mail: wohlmuth@ma.tum.de